

DOCUMENT RESUME

ED 218 283

TM 820 226

AUTHOR Berger, Carl F.
TITLE Trend Analysis Using Microcomputers.
PUB DATE Apr 81
NOTE 59p.; Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (April, 1981).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Analysis of Variance; *Computer Programs; *Data Processing; Elementary Secondary Education; *Microcomputers; Science Education; Statistical Analysis; *Trend Analysis
IDENTIFIERS *Apple II

ABSTRACT

A trend analysis statistical package and additional programs for the Apple microcomputer are presented. They illustrate strategies of data analysis suitable to the graphics and processing capabilities of the microcomputer. The programs analyze data sets using examples of: (1) analysis of variance with multiple linear regression; (2) exponential and higher order regression; (3) orthogonal tests of trend; and (4) the use of dummy variables to handle categorical variables. Microcomputer benefits and limitations are highlighted with the exploratory analysis techniques. An example of a traditional analysis of variance test is presented which illustrates the microcomputer data entry procedure. Graphic displays of program statements and diagrams are included with an appendix of sample programs. (CM)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

FEB 8 1982

ED218283

TREND ANALYSIS USING MICROCOMPUTERS

Carl F. Berger

The University of Michigan

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Berger, Carl F.

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)"

Presented at the
Annual Meeting of the
National Association for Research
in Science Teaching

April 1981

-TREND STATISTICAL PACKAGE-

STATISTICAL PROGRAMS

FOR THE APPLE BY

CARL F. BERGER
UNIVERSITY OF MICHIGAN.

DISC VOLUME 254

```

* A 012 ANOVAKP
* B 014 AS CHR GEN
* B 006 CHARACTER TABLE
* B 014 COUNT
* A 009 EDIT
* A 013 EXPOREGRESS
* A 016 FNXGRAPH/PADDLES
* A 012 GRAPH RESULTS
* A 002 HELLO
* B 003 HI-RES CHARACTER GENERATOR
* A 011 HLO
* B 002 LOMEM:
* A 012 MLRKP
* A 018 MLRTWOWAYKP
* A 015 NTH-COR
* A 016 NTH-CORKP
* T 002 THLO
* A 004 TITLE
* A 019 TRENDKP
* A 014 TWOWAYKP
  A 002 DATA
  A 013 EXPOREGRESSKP
* R 019 APA
* A 003 LOADAPA
* B 003 RBOOT
* B 005 RLOAD

```

Looking for relationships is a natural role of science education. More and more it has become important not only to test for differences in groups but also to analyze data for trends. Finding relationships in data sets gives more information than traditional tests of mean differences or traditional ANOVA tests or even MANOVA tests. The use of techniques such as Multiple Linear Regression, Higher Order Correlation, and Exponential Regression can often indicate relationships too easily overlooked in traditional research.

Trend analysis has not been used widely, partly because of the need for extra information added to existing data in order to search for trends. Further, the reliance on numerical values rather than graphs to indicate trend has not provided the necessary comfort needed to explore data.

The microcomputer has capabilities which overcome these drawbacks. First, the graphics mode of the television screen offers greater possibilities in looking for trends. Second, the ability to easily adapt data sets encourages the search for trends. Third, the microcomputer can be used to compare several kinds of analyses in order to get a real feel for the structure and implications inherent in the data.

In this session, the microcomputer will be used to:

- 1.. Compare alternate strategies of data analysis.
2. Graphically show results of tests for trend.

3. Indicate the powerful application of such machines for reasonable size data sets.

Analysis of data sets will be examined using a comparison of:

- analysis of variance with Multiple Linear Regression;
- the use of Exponential Regression and Higher Order Regression;
- the use of orthogonal tests of trend; and
- the use of dummy variables to handle categorical variables.

Throughout the session the use of exploratory analysis techniques will be emphasized. The limitations as well as the benefits of the use of microcomputers will be highlighted.

Sample programs are available in the appendix; printouts of t.v. screen data are included. The examples used are from Kerlinger and Pedhazur¹, with specific page references included.

The first analysis is a traditional ANOVA on a set of data from page 106 in Kerlinger and Pedhazur. The analysis is straightforward but illustrates an important data entry procedure on the microcomputer. Data should be included within programs as program steps. The entry of data as input on demand from the program means that any errors, or modifications require the entry of the entire set from start to finish. If data is included as part of data statements with the program, modifications can easily be made, errors corrected, and more data added with little difficulty.

So as shown in the following example, the data is for three groups and is listed within the program as lines 500-999. (The shadowed lines are responses to computer prompts.) Note that the data for three groups is indicated by line 500 data 3. The data follows on lines 501, 502, 503.

To modify data, merely retype a line such as 502 data 7, 8, 9, 10, 10, 999. Any subsequent run would reflect this modification. To save modified data, type SAVE (or CSAVE), and the name of the program (in quotes or not depending on the brand of microcomputer you are using).

Some comments on the analysis. In addition to the standard F value, the values of ETA square, Omega square and the probability level are included.

IRUN

ANALYSIS OF VARIANCE

STATISTICAL PROGRAMS

FOR THE APPLE BY

CARL F. BERGER
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO START

ANALYSIS OF VARIANCE

THIS PROGRAM CALCULATES THE F VALUE
USED TO FIND IF THERE IS A DIFFERENCE
IN THE MEANS OF GROUPS OF DATA.

NEED INSTRUCTIONS?

?YES

TYPE IN YOUR DATA STARTING ON LINE 500. TYPE 500 DATA K

501 DATA A1,A2,A3,...999

502 DATA B1,B2,B3,...999 ETC. WHERE K IS THE NUMBER OF GROUPS AND A, B, ARE
ATAPOINTS

ENTER ONLY 40 NUMBERS PER LINE AND LIST K AS THE FIRST NUMBER ONLY

999 STOPS THE DATA SET FOR EACH GROUP.

HERE'S AN EXAMPLE

500 DATA 3

501 DATA 4,5,6,7,8,999

502 DATA 7,8,9,10,11,999

503 DATA 1,2,3,4,5,999

PRESS <RETURN> TO START THE NUMBER OF GROUPS IS 3

N	MEAN	STANDARD DEV
5	6	1.58113883
5	9	1.58113882
5	3	1.58113883

UNIVARIATE 1-WAY ANOVA

SOURCE	SUM OF SQ.	D.F.	MEAN SQ.
BETWEEN	90.0000001	2	45.0000001
WITHIN	29.9999998	12	2.49999998
TOTAL	120	14	

THE F VALUE IS 18.0000002

WITH 2 AND 12 DEGREES OF FREEDOM

ETA SQUARE= .750000001

OMEGA SQUARE= .693877553

PROBABILITY LEVEL IS 4E-04

The subroutines to calculate significance levels and some other subroutines are from Some Basic Statistical Programs.²

To introduce an alternate form of analysis that encourages a test for trend, the same data was analyzed using Multiple Linear Regression (MLR) with the following results.

Note similar F values, an R^2 equivalent to ETA^2 and a standard error of the estimate equivalent to the standard deviations of each value (sort of an average standard deviation!).

The means of the data sets are produced by the use of the prediction equation

$$y = 6(X_1) + 3(X_2) + 6$$

X_1 and X_2 are dummy variables to account for the groups.

Two dummy variables are needed for three groups: a code of 1 and 0 for Group 1, 0 and 1 for Group 2, and 0 and 0 for Group 3. fully cover the relationships.

The use of MLR is preferred by many researchers because of the calculation of the standard error of estimate and the relaxation of problems due to unequal N's.


```

- LINEAR REGRESSION ANALYSIS-
  STATISTICAL PROGRAMS

  FOR THE APPLE BY

  CARL F. BERGER
  UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO START

```

MULTIPLE LINEAR REGRESSION

NEED INSTRUCTIONS?YES
TYPE IN DATA AT LINE 900,
DATA,NUMBER OF INDEPENDENT VAR,SAMPLE SIZE,
EACH INDEPENDENT VARIABLE AND EACH DEPENDENT VARIABLE
HERE IS LINE 900 AS A SAMPLE

```
900 DATA 2,15,1,0,4,1,0,5,1,0,6,
1,0,7,1,0,8,0,1,7,0,1,8,0,1,
9,0,1,10,0,1,11,0,0,1,0,0,2,
0,0,3,0,0,4,0,0,5
```

PRESS <RETURN> TO STARTSAMPLE SIZE IS 15
NUMBER OF INDEPENDENT VARIABLES ARE 2

1 0 4
1 0 5
1 0 6
1 0 7
1 0 8
0 1 7
0 1 8
0 1 9
0 1 10
0 1 11
0 0 1
0 0 2
0 0 3
0 0 4
0 0 5

EQUATION COEFFICIENTS:

```

CONSTANT: 3
VARIABLE(1): 3
VARIABLE(2): 6

```

COEFFICIENT OF DETERMINATION(R-SQR)=.75

COEFFICIENT OF MULTIPLE CORRELATION = .866025404

STANDARD ERROR OF ESTIMATE .158113883

F VALUE = 18.0000001
WITH 2 AND 12 DEGREES OF FREEDOM

INTERPOLATION: (ENTER 9999 TO END PROGRAM)

VARIABLE1?1

VARIABLE2?0

DEPENDENT VARIABLE =6.00000001

VARIABLE1?0

VARIABLE2?1

DEPENDENT VARIABLE =9

VARIABLE1?0

VARIABLE2?0

DEPENDENT VARIABLE =3

VARIABLE1?9999

Two way analysis of variance is another procedure that can be analyzed on the microcomputer.

Here the data are taken from page 156 of Kerlinger and Pedhazur.

TEACHING METHODS

	Ausabel	Piaget	Skinner
Biology	16 14	20 16	10 14
Chemistry	12 10	17 13	7 7
Physics	7 7	10 8	6 4

A traditional two way analysis often looks like this.

PR#0
JRUN

-TWO WAY ANALYSIS OF VARIANCE-

STATISTICAL PROGRAMS

- FOR THE APPLE BY

CARL F. BERGER
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO CONTINUE

TWO WAY ANALYSIS OF VARIANCE
NEED INSTRUCTIONS?

?YES

INSERT DATA ON LINE 1080 AS 1080 DATA C,R,X11,...,999,X12,...,999 WHERE C IS THE
NUMBER OF COLUMNS AND R IS THE NUMBER OF ROWS

1080 DATA 3,3

1090 DATA 16,14,999,12,10,999,7,
7,999

1100 DATA 20,16,999,17,13,999,10,
8,999

1120 DATA 10,14,999,7,7,999,4,6,
999

PRESS <RETURN> TO CONTINUE THE NUMBER OF GROUPS IS. 9

N	MEAN	STANDARD DEV
2	15	1.41421356
2	11	1.41421354
2	7	1.88740797E-04
2	18	2.82842707
2	15	2.82842713
2	9	1.41421355
2	12	2.82842712
2	7	1.88740797E-04
2	5	1.41421356

UNIVARIATE 1-WAY ANOVA

SOURCE	SUM OF SQ.	D.F.	MEAN SQ.
--------	------------	------	----------

BETWEEN	308	8	38.5
---------	-----	---	------

WITHIN	31.9999991	9	3.55555545
--------	------------	---	------------

TOTAL	339.999999	17
-------	------------	----

THE F VALUE IS 10.8281253

WITH 8 AND 9 DEGREES OF FREEDOM

ETA SQUARE= .905882356

PRESS <RETURN> TO CONTINUE TWO WAY ANALYSIS OF VARIANCE

SOURCE OF VARIATION	SUM OF SQUARES	DEGREES OF FREEDOM
------------------------	-------------------	-----------------------

ROWS	192	2
COLUMNS	108	2
INTERACTION	8.00000024	4
RESIDUAL	31.99999991	9
TOTAL	339.999999	17

VARIANCE ESTIMATE	F VALUE	ETA SQUARE LEVEL
96	27	.318
54	15.188	.565
2.000000006	.563	.024

Notice a very high ETA square of .906 for a univariate ANOVA and the sum of the ETA square levels of $.318 + .565 + .024 = .907$. Thus the explained variance in the univariate analysis hides the explained variances of the teaching methods, subject matter, and interaction. (As might be expected, the greatest difference is in subject matter, then teaching method, and finally no interaction of teaching method and subject matter.)

Two way analysis of variance can also be analyzed using techniques of Multiple Linear Regression.

The example here is from page 174 of Kerlinger and Pedhazur. First the data must be recast into a form necessary for MLR. Here we will use effect coding.

Dummy Variable

1	2	3	4	5	6	7	8	
1	0	1	0	1	0	0	0	16/14
0	1	1	0	0	0	1	0	12/10
-1	-1	1	0	-1	0	-1	0	7/7
1	0	0	1	0	1	0	0	20/16
0	-1	0	1	0	0	0	1	17/13
-1	-1	0	1	0	-1	0	-1	10/8
1	0	-1	-1	-1	-1	0	0	10/14
0	1	-1	-1	0	0	-1	-1	7/7
-1	-1	-1	-1	1	1	1	1	4/6

	1	2	3
1	16 14	12 10	7 7
2	20 16	17 13	10 8
3	10 14	7 7	4 6

IMPORTANT INTERACTION

.00

1	2	3
1	0	-1
1	0	-1
1	0	-1

Columns

.001

4	5	6
0	0	0
1	1	1
-1	-1	-1

Rows

.166

8	9	10
0	0	0
0	1	-1
0	-1	1

Interaction

OTHER INTERACTIONS OF INTEREST

.005

2	3	4
0	1	-1
0	1	-1
0	1	-1

.021

3	4	5
1	1	1
0	0	0
-1	-1	-1

NOT MUCH INTEREST

5

1	0	-1
0	0	0
-1	0	1

.716

6

0	0	0
.1	0	-1
-1	0	1

.522

7

0	1	-1
0	0	0
0	-1	1

.522

Again, ETA^2 and R^2 are equivalent in a one way design. In the two way analysis, the rows and columns are represented by two dummy variables. Here a field dummy variable is used, in which 0 and -1 are used so a complete orthogonal effect is obtained.

Using single order correlations, each individual combination can be examined using MLR.

The row ETA^2 equals the 4th single order R^2 .

The column ETA^2 equals the 1st single order R^2 .

The interaction ETA^2 equals the 8th single order R^2 .

In addition, other correlations can be examined for significance such as the second and third single order R^2 values.

Casting the dummy variables in their associated rows and columns allows us to see the relationships. Further, we can start our search for trend.

JNRUN

-MULTIPLE LINEAR REGRESSION-
- TWO WAY ANALYSIS ADDED -
STATISTICAL PROGRAMS

FOR THE APPLE BY-

CARL F. BERGER
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO CONTINUE

MULTIPLE LINEAR REGRESSION

NEED INSTRUCTIONS?YES

TYPE IN DATA AT LINE 900,

DATA,NUMBER OF INDEPENDENT VAR,SAMPLE SIZE,

EACH INDEPENDENT VAR,AND EACHDEPENDENT VARIABLE

HERE IS LINE 900 AS A SAMPLE

900 DATA 8,18,1,0,1,0,1,0,0,0,16
1,0,1,0,1,0,0,0,14

PRESS <RETURN> TO CONTINUE
SAMPLE SIZE IS 18
NUMBER OF INDEPENDENT VARIABLES ARE 8

1 0 1 0 1 0 0 0 16
1 0 1 0 1 0 0 0 14
0 1 1 0 0 0 1 0 12
0 1 1 0 0 0 1 0 10
-1 -1 1 0 -1 0 -1 0 7
-1 -1 1 0 -1 0 -1 0 7
1 0 0 1 0 1 0 0 20
1 0 0 1 0 1 0 0 16
0 1 0 1 0 0 0 1 17
0 1 0 1 0 0 0 1 13
-1 -1 0 1 0 -1 0 -1 10
-1 -1 0 1 0 -1 0 -1 8
1 0 -1 -1 -1 -1 0 0 10
1 0 -1 -1 -1 -1 0 0 14
0 1 -1 -1 0 0 -1 -1 7
0 1 -1 -1 0 0 -1 -1 7
-1 -1 -1 -1 1 1 1 1 4
-1 -1 -1 -1 1 1 1 1 6

THANKS FOR THE DATA. NOW LETS SEE...

EQUATION COEFFICIENTS:

CONSTANT: 11
VARIABLE(1): 4
VARIABLE(2): 0
VARIABLE(3): 0
VARIABLE(4): 3
VARIABLE(5): 0
VARIABLE(6): 0
VARIABLE(7): 0

VARIABLE(8): 1

R-SQUARE=.905882355

MULTIPLE CORRELATION =.951778522

STANDARD ERROR OF ESTIMATE 1.88561806

F VALUE = 10.8281252

WITH 8 AND 9 DEGREES OF FREEDOM

SIGNIFICANCE LEVEL IS 1.3E-03

SINGLE ORDER CORRELATIONS?

PRESS <RETURN> TO CONTINUEVAR# R-SQR F SIG LEVEL

1	.565	.54	0
2	.141	13.5	5E-03
3	.079	7.594	.021
4	.318	30.375	1E-03
5	1E-03	.141	.716
6	6E-03	.563	.522
7	6E-03	.563	.522
8	.024	2.25	.166

PRESS <RETURN> TO CONTINUEINTERCORRELATIONS

1 2	.5
1 3	0
1 4	0
1 5	0
1 6	0
1 7	0
1 8	0
1 9	.751
2 3	0
2 4	0
2 5	0
2 6	0
2 7	0
2 8	0
2 9	.376
3 4	.5
3 5	0
3 6	0
3 7	0
3 8	0
3 9	.282
4 5	0
4 6	0
4 7	0
4 8	0
4 9	.564
5 6	.5
5 7	.5
5 8	.25
5 9	.038
6 7	.25
6 8	.5
6 9	.077
7 8	.5
7 9	.077
8 9	.153

PAGES 175-180 FOR PROPORTIONS OF
VARIANCE

Now we turn to data which lend themselves more directly to a search for trend.

First a look at data which show definite trend. The analysis is Nth - correlation analysis. The data set is from page 210 of Kerlinger and Pedhazur.

The first step is a plot of the data which has the following values.

X	Y	X	Y	X	Y
2	4	6	13	10	18
2	6	6	14	10	19
2	5	6	15	10	20
4	7	8	16	12	19
4	10	8	17	12	20
4	10	9	21	12	21

J

JRUN

-FITTING HIGHER ORDER PLOTS-

CONTAINS SOME COMMON BASIC SUB
PROGRAMS BY

FOR THE APPLE BY
CARL F. BERGER
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO START

CORRELATION PLOTTING
DO YOU NEED INSTRUCTIONS?
YES

```
490 REM *** INSTRUCTIONS ***
500 REM TYPE IN DATA AS 510 D
    DATA X1,Y1,...999,999
502 REM WHERE X,Y ARE X VALUES
    Y VALUES.
504 REM USE 999,999,999 TO STOP

510 DATA 2,4,2,6,2,5,4,7,4,10
    ,4,10,6,13, 6,14,6,15,8,16,8
    ,17,8,21,10,18,10,19,10,20,1
    2,19,12,20,12,21
999 DATA 999,999,999
```

JCONT

21

21

12

DEGREE OF THE EQUATION DESIRED

21

JRUN

-FITTING HIGHER ORDER PLOTS-

CONTAINS SOME COMMON BASIC SUB
PROGRAMS BYFOR THE APPLE BY
CARL F. BERGER
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO START

CORRELATION PLOTTING
DO YOU NEED INSTRUCTIONS?
YES

```

490 REM *** INSTRUCTIONS ***
500 REM TYPE IN DATA AS 510 D
    ATA X1,Y1,...999,999
502 REM WHERE X,Y ARE X VALUES
    , Y VALUES.
504 REM USE 999,999,999 TO STOP

510 DATA 2,4,2,6,2,5,4,7,4,10
    ,4,10,6,13, 6,14,6,15,8,16,8
    ,17,8,21,10,18,10,19,10,20,1
    2,19,12,20,12,21
999 DATA 999,999,999

```

JCONT

DEGREE OF THE EQUATION DESIRED1
THIS MAY TAKE A LITTLE TIME....HMM..
STILL THINKING!
STILL THINKING!
STILL THINKING!
STILL THINKING!

CONSTANT= 3.2666667
1 DEGREE COEFFICIENT= 1.55714286

COEFFICIENT OF DETERMINATION (R SQR)= .88323628
COEFFICIENT OF CORRELATION= .939806512
STANDARD ERROR OF THE ESTIMATE= 2.05113207

CARE TO LOOK AT THE BEST FIT? N
ANOTHER FIT? Y

-21-

23

ANOTHER FIT?

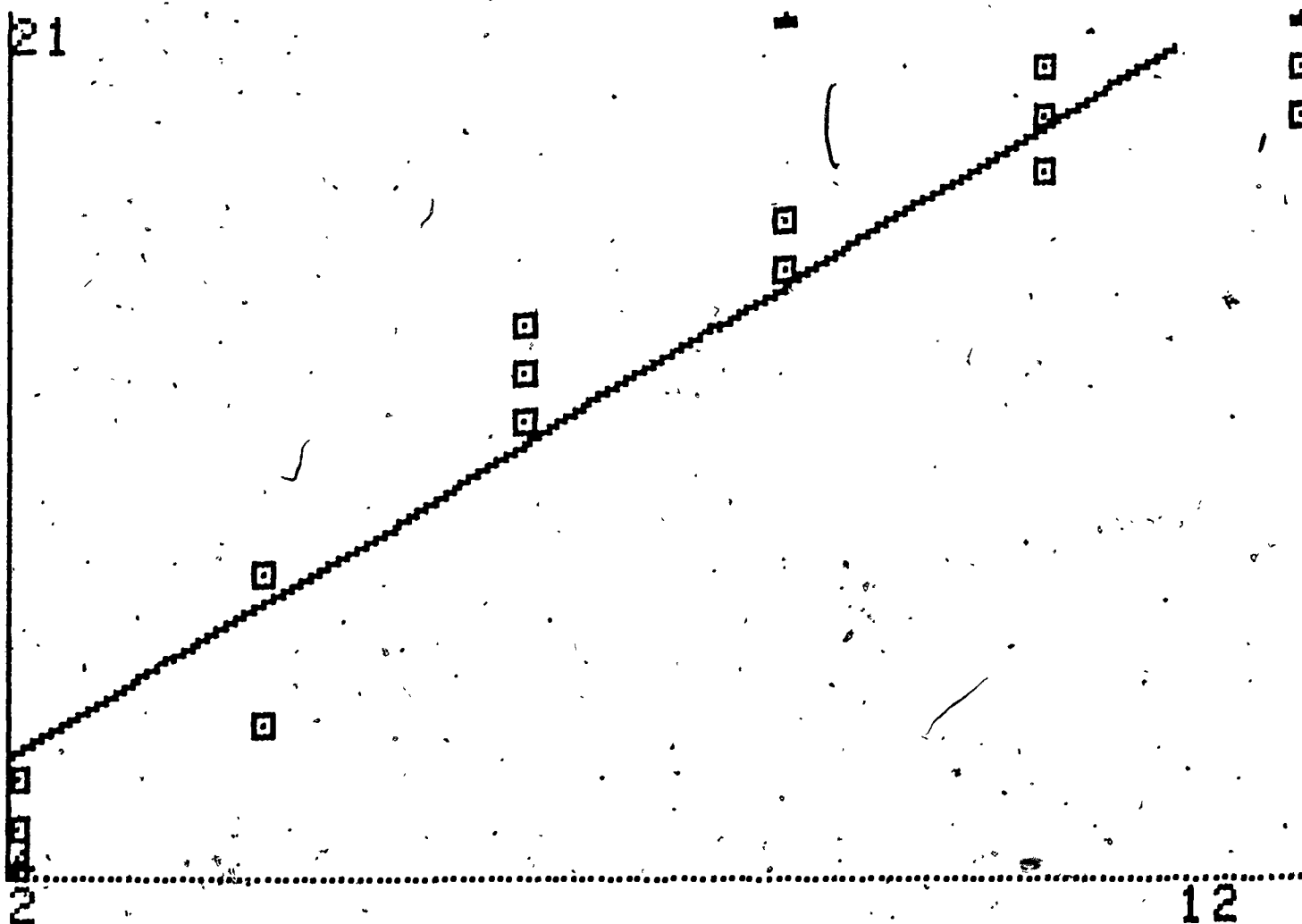
21

2

12

24

ANOTHER FIT? YES



?REDIM'D ARRAY ERROR IN 10310
JRUN

-FITTING HIGHER ORDER PLOTS-

CONTAINS SOME COMMON BASIC SUB
PROGRAMS BY

FOR THE APPLE BY
CARL F. BERGER
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO START

CORRELATION PLOTTING
DO YOU NEED INSTRUCTIONS?
?N

DEGREE OF THE EQUATION DESIRED?2
THIS MAY TAKE A LITTLE TIME...HMM..

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

CONSTANT= -1.90000019

1 DEGREE COEFFICIENT= 3.49464293

2 DEGREE COEFFICIENT= -138392862

COEFFICIENT OF DETERMINATION (R SQR)= .942770418

COEFFICIENT OF CORRELATION= .970963655

STANDARD ERROR OF THE ESTIMATE= 1.48307909

CARE TO LOOK AT THE BEST FIT? N
ANOTHER FIT? N

PRESS <RETURN> TO START

CORRELATION PLOTTING

DO YOU NEED INSTRUCTIONS?

YN

DEGREE OF THE EQUATION DESIRED?

THIS MAY TAKE A LITTLE TIME....HMM...

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

CONSTANT= .666666589

1 DEGREE COEFFICIENT= 1.88029105

2 DEGREE COEFFICIENT= .128948252

3 DEGREE COEFFICIENT= -.0127314812

COEFFICIENT OF DETERMINATION (R²)= .946268567

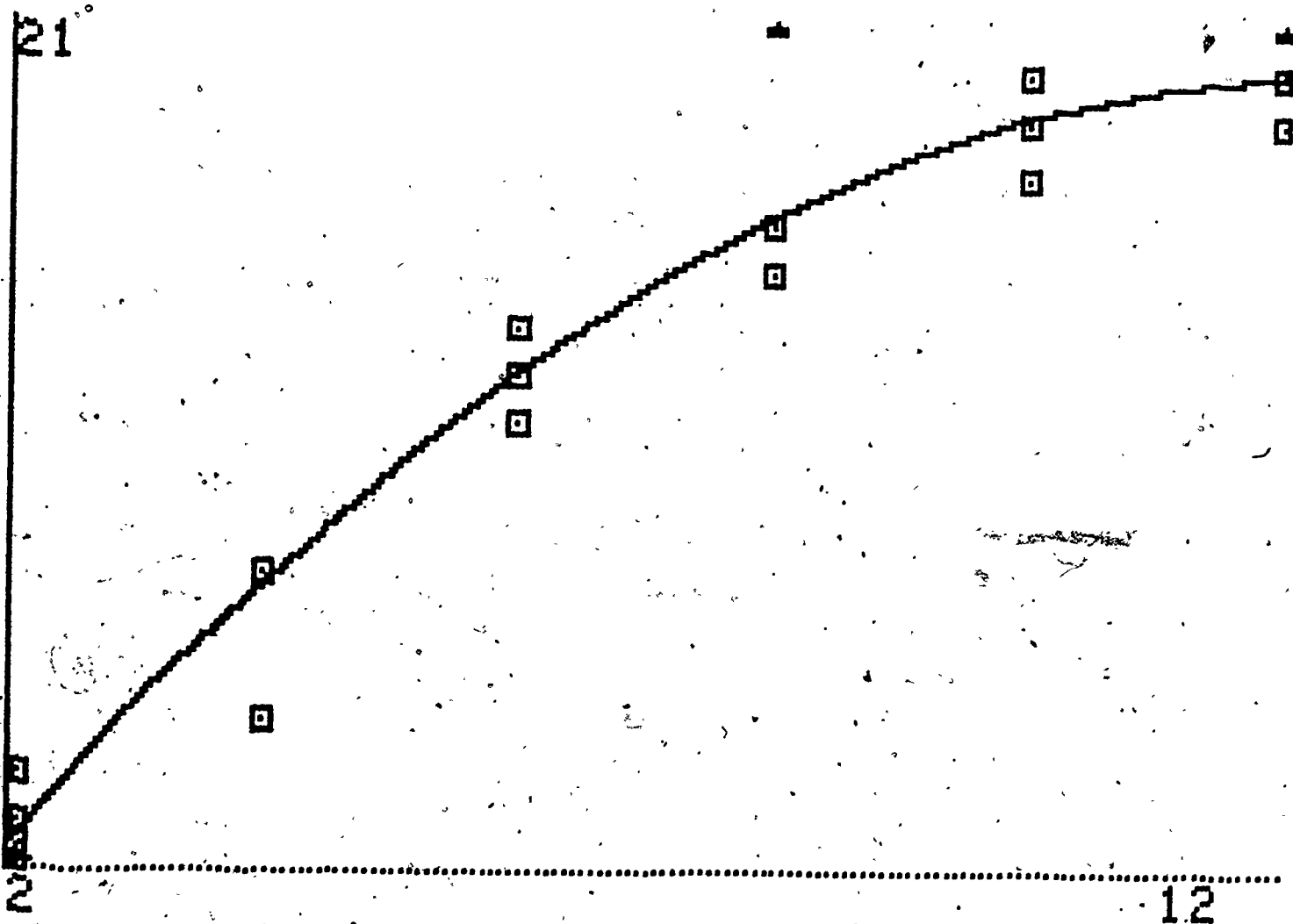
COEFFICIENT OF CORRELATION= .972763366

STANDARD ERROR OF THE ESTIMATE= 1.4874756

CARE TO LOOK AT THE BEST FIT? N

ANOTHER FIT? N

21



12

ANOTHER FIT? YES

27

CORRELATION PLOTTING

DO YOU NEED INSTRUCTIONS?

?N

DEGREE OF THE EQUATION DESIRED4

THIS MAY TAKE A LITTLE TIME....HHH..

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

STILL THINKING!

CONSTANT= 8.16669438

1 DEGREE COEFFICIENT= -4.57806223

2 DEGREE COEFFICIENT= 1.8663239

3 DEGREE COEFFICIENT= -.195023529

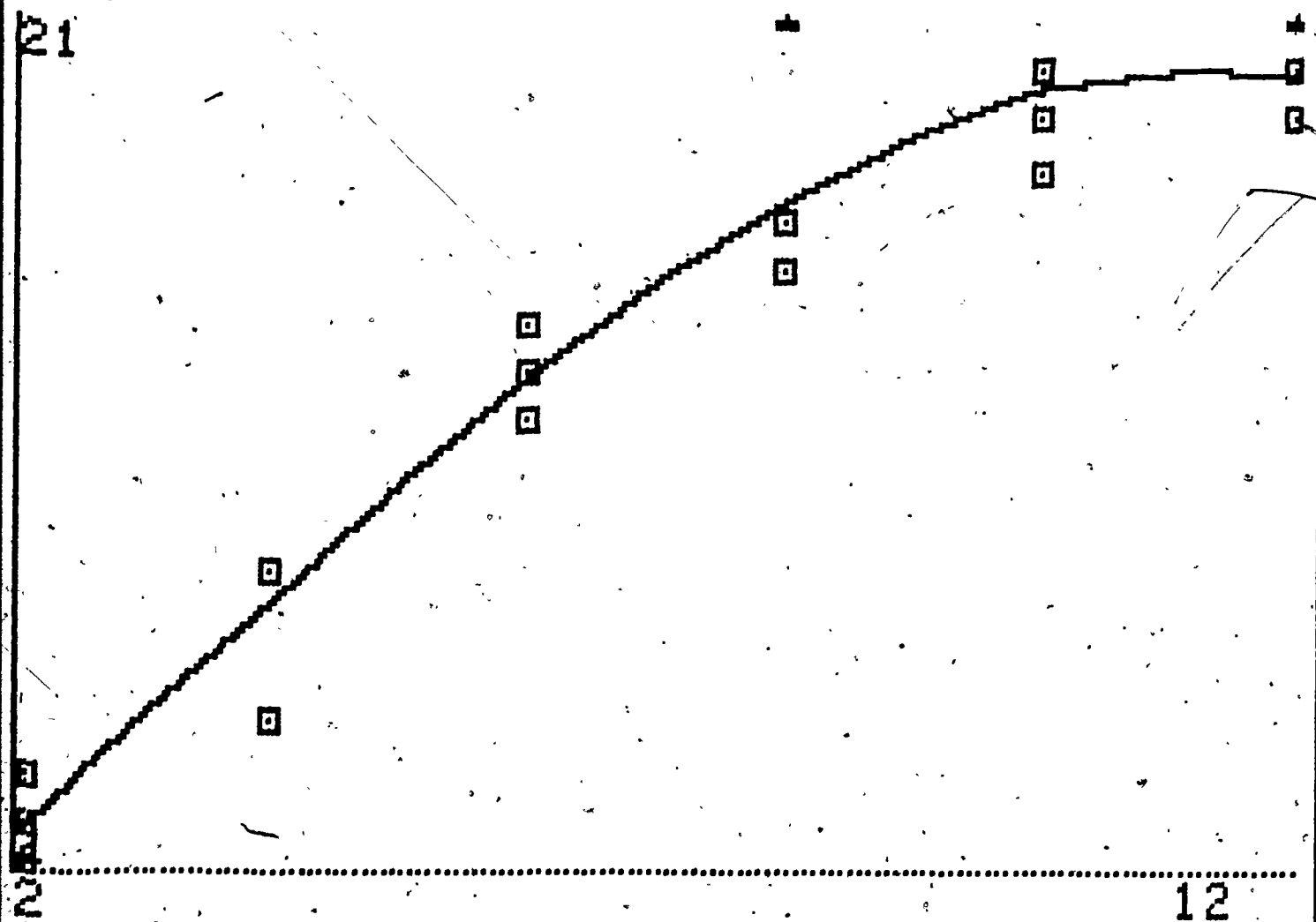
4 DEGREE COEFFICIENT= 6.51042802E-03

COEFFICIENT OF DETERMINATION (R SQR)= .950915275

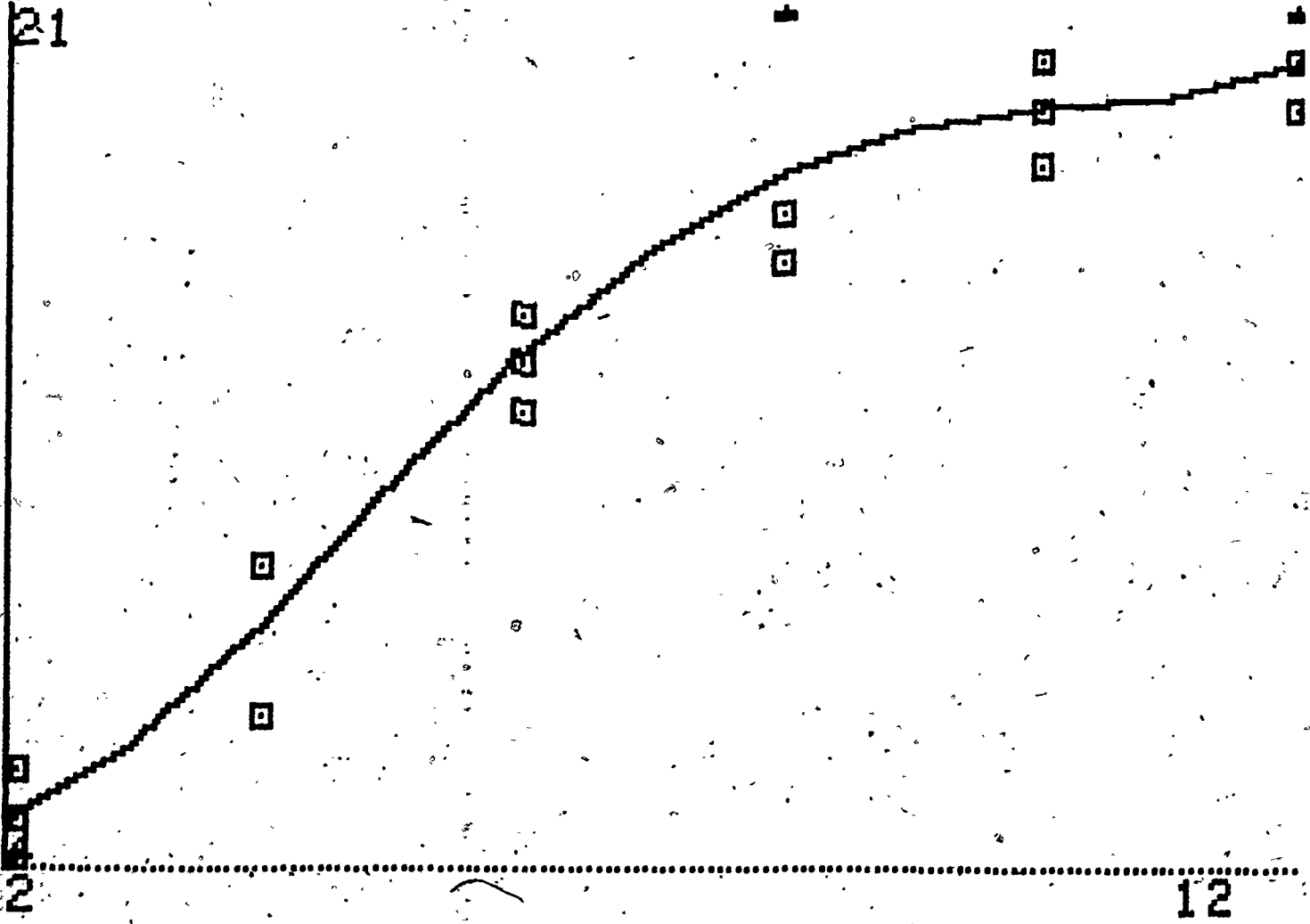
COEFFICIENT OF CORRELATION= .975148848

STANDARD ERROR OF THE ESTIMATE= 1.47537072

CARE TO LOOK AT THE BEST FIT? Y

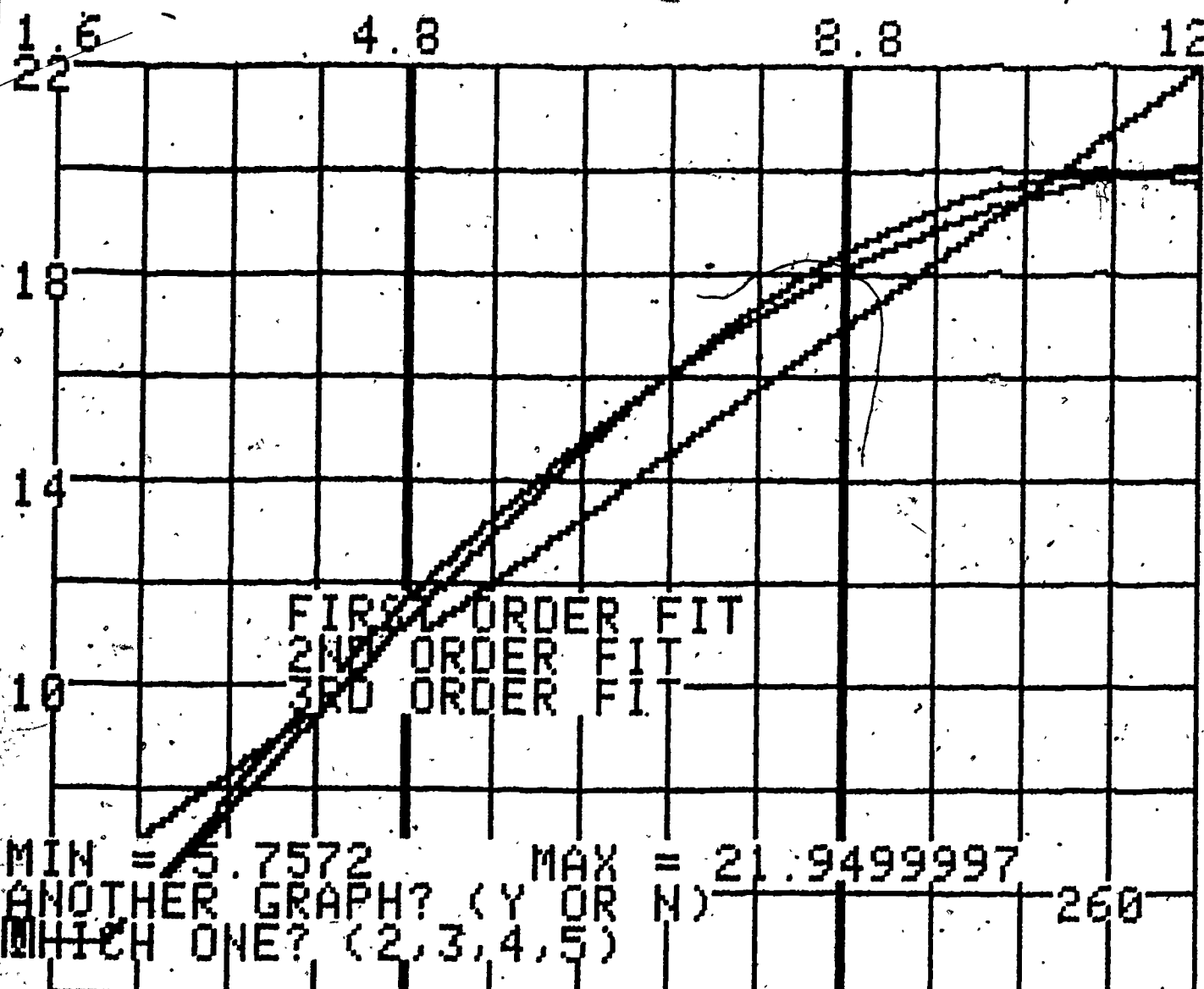


21



ANOTHER FIT? NO

30



Exponential Regression

In addition to higher order polynomial fits (another name for the analysis just completed), we can examine an exponential fit of the form

$$y = e^A x e^{BXA}$$

For some rare science education data such fits may work. The example shown is from the findings of students measuring the force and distance between attracting magnets.

The data are from six separate student teams.

Distance between magnets	Force in washers to release
0	11, 20, 18, 19, 19, 19
1	9, 12, 12, 10, 10, 13
2	5, 6, 6, 7, 8, 8, 9
3	4, 5, 5, 6, 6, 7
4	3, 3, 4, 4, 5, 6
5	2, 2, 2, 4, 4, 4, 5

The computer printout and values are listed on the following page.

JRUN

- EXPONENTIAL REGRESSION -

STATISTICAL PROGRAMS

FOR THE APPLE BY

CARL F. BERGER.
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO START

CORRELATION PLOTTING
DO YOU NEED INSTRUCTIONS?
YES

```
490 REM *** INSTRUCTIONS ***
500 REM TYPE IN DATA AS 510 DAT
    A X1,Y1,F1,...,999,999,999
502 REM WHERE X,Y,F ARE X VALUE
    S, Y VALUES, FREQUENCIES.
504 REM USE 999,999,999 TO STOP

510 DATA 0,11,1,0,20,1,0,18,1,0,
    19,3
520 DATA 1,9,1,1,12,2,1,10,2,1,1
    3,0
530 DATA 2,5,1,2,6,1,2,7,1,2,8,2
    ,2,9,1
540 DATA 3,4,1,3,5,2,3,6,2,3,7,1

550 DATA 4,3,2,4,4,2,4,5,1,4,6,1

560 DATA 5,2,3,5,4,2,5,5,1
999 DATA 999,999,999
```

EXPONENTIAL REGRESSION

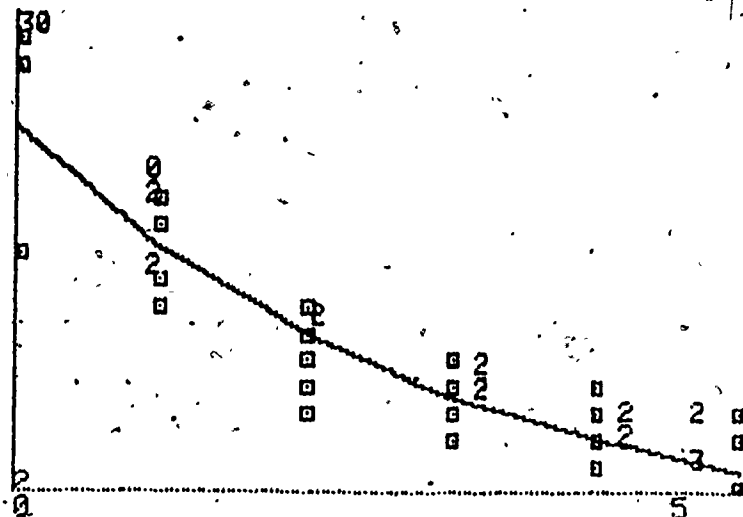
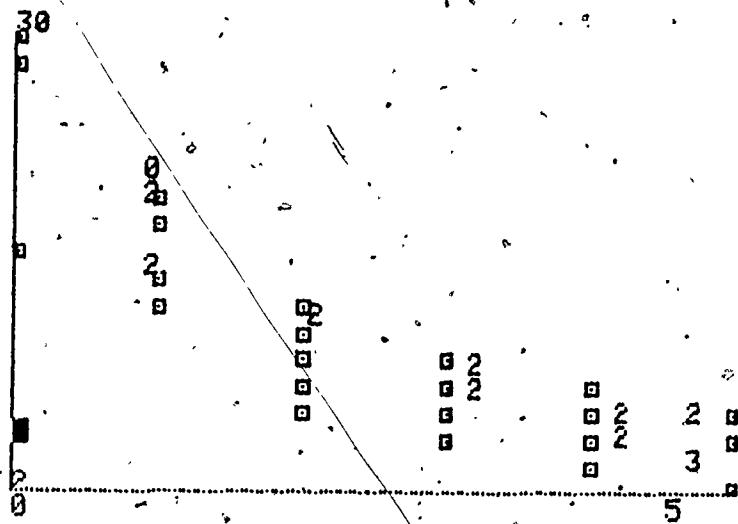
A= 15.7742992
B= -.346410937

COEFFICIENT OF DETERMINATION (R²)= .849172246
COEFFICIENT OF CORRELATION= .921505424
STANDARD ERROR OF THE ESTIMATE= .25656034

CARE TO LOOK AT THE BEST FIT? Y

20
2
0

34



ANOTHER FIT? N

35

Tests for Trend Using Orthogonal
and/or Person Vectors

As a final example from Kerlinger and Pedhazur, we use the data set from pages 218 and 219 in a section entitled Trend Analyses with Repeated Measures.

Here not only are there orthogonal variables but also a test for repeated measures--that is, a measure for each person so we can note if the influence of individual differences will make a significant contribution to the explanation of variation of scores.

Note how the data are entered. A small subroutine will generate the dummy variables for the replicated samples and the orthogonal variables are entered as separate data statements starting on line 160.

Having the data in this form saves the massive data entry problems associated with this kind of analysis technique.

JRUN

- REGRESSION ANALYSIS WITH -
- REPLICATION OF PERSONS -
STATISTICAL PROGRAMS

FOR THE APPLE BY

CARL F. BERGER
UNIVERSITY OF MICHIGAN

PRESS <RETURN> TO CONTINUE
REGRESSION ANALYSIS

NEED INSTRUCTIONS? YES

INSERT THE NUMBER OF REPLICATES AT LINE 20

THEN TIMES OR PROBLEMS AT LINE 30

THE CONTRASTS FOR TREND AT LINES 40 AND

AT LINE 160

AND TYPE INDEPENDENT VARIABLES STARTING AT LINE 1550

20 N = 3

25 REM REPLICATES/SUBJECTS

30 P = 6

35 REM TIMES/PROBLEMS

40 TT = 5

45 REM CONTRASTS FOR TREND

160 DATA -5,5,-5,1,-1,-3,-1,7,-3
5

1550 DATA 4,6,5,7,10,10,13,14,15
16,17,21,18,19,20,19,20,21

JRUNPR#1

1

JCONT

SAMPLE SIZE IS 18

-5 5 -5 1 -1 1 0 4

-5 5 -5 1 -1 0 1 6

-5 5 -5 1 -1 0 0 5

-3 -1 7 -3 5 1 0 7

-3 -1 7 -3 5 0 1 10

-3 -1 7 -3 5 0 0 10

-1 -4 4 2 -10 1 0 13

-1 -4 4 2 -10 0 1 14

-1 -4 4 2 -10 0 0 15

1 -4 -4 2 10 1 0 16

1 -4 -4 2 10 0 1 17

1 -4 -4 2 10 0 0 21

3 -1 -7 -3 -5 1 0 18

3 -1 -7 -3 -5 0 1 19

3 -1 -7 -3 -5 0 0 20

As the output indicates, the single order correlations can account for each variable separately. But the total shared variation of the single correlations squared do not equal the total R square finally printed out, .9796 versus .9844. The only way to match more closely the total R^2 is to use effect codes and better yet full orthogonal dummy variables.¹

Nevertheless, the variation can still be isolated and is due to:

- 1 - a linear component
 - 2 - a quadratic component
 - 3 - not a cubic component
- and some significant individual correlation in variable 6.

5 5 5 1 1 1 0 19
5 5 5 1 1 0 1 20
5 5 5 1 1 0 0 21

EQUATION COEFFICIENTS:

CONSTANT: 15.3333333
VARIABLE(1): 1.55714286
VARIABLE(2): -.369047619
VARIABLE(3): -.0611111111
VARIABLE(4): .178571429
VARIABLE(5): .0198412698
VARIABLE(6): -2.5
VARIABLE(7): -1

PRESS <RETURN> TO CONTINUE

COEFFICIENT OF DETERMINATION(R-SQR)=.984388554

COEFFICIENT OF MULTIPLE CORRELATION =.992163573

STANDARD ERROR OF ESTIMATE .94868321

F VALUE = 90.0793819

WITH 7 AND 10 DEGREES OF FREEDOM

SIGNIFICANCE LEVEL IS 0

SINGLE ORDER CORRELATIONS?YES

VAR#	R-SQR	F	SIG LEVEL
1	.883	56.776	0
2	.06	38.13	2E-04
3	3E-03	2.24	.1629
4	5E-03	2.98	.1126
5	1E-03	.58	.5833
6	.028	17.78	2.1E-03
7	0	.61	.6144

TOTAL R SQR IS .979618383

.LIST1550-

1550 DATA 4,6,5,7,10,10,13,14,15
16,17,21,18,19,20,19,20,21
1560 DATA 101.6,89.6,85.3,106.2,
112.8,84,73.4,68,108,120
1570 DATA 61.6,91.1,78.8,91.4,94
80.2,72.6,64,63.2,94

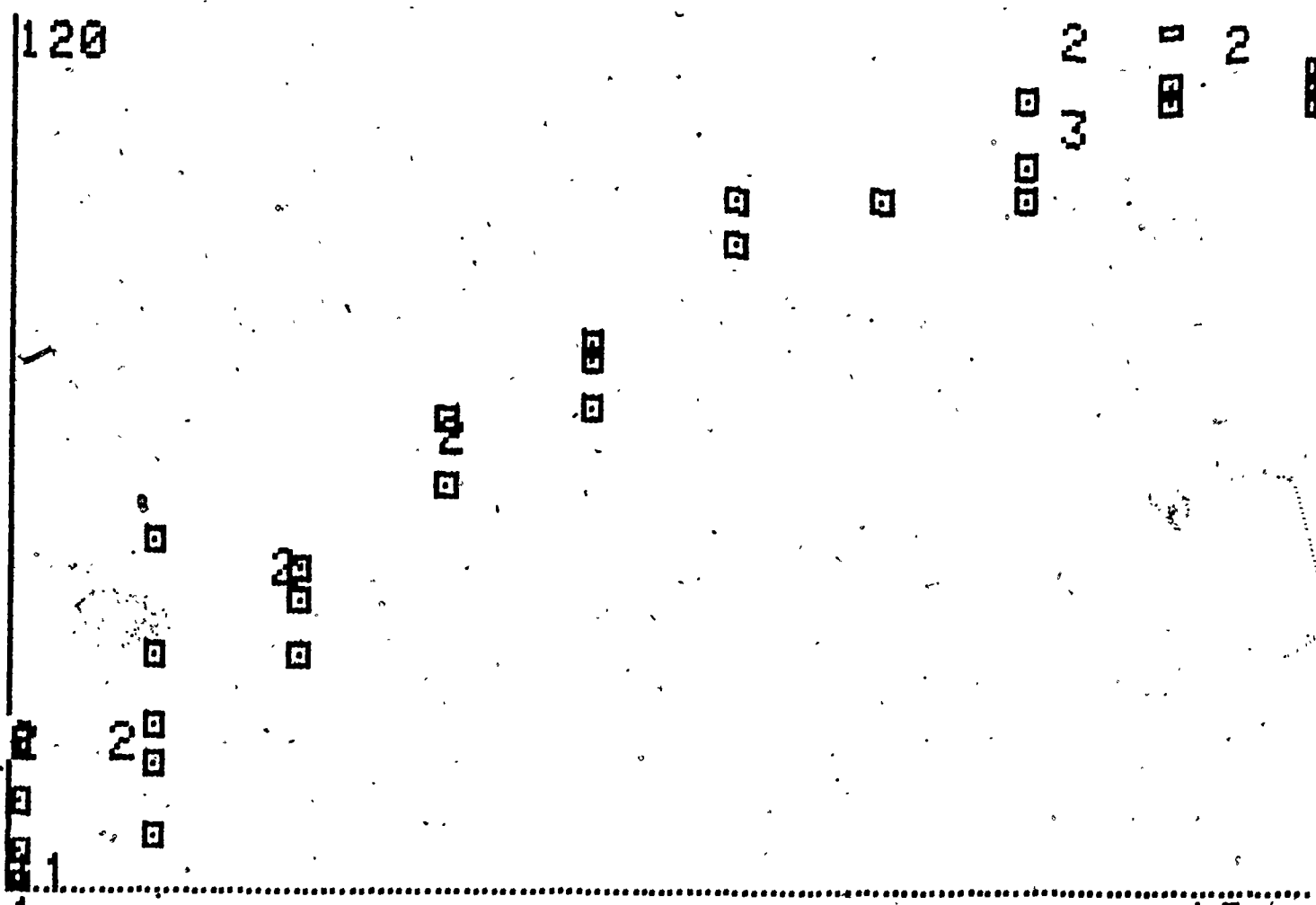
Some sample problems to test the ideas presented:

CONSTANT= 15.7927155
1 DEGREE COEFFICIENT= 10.5436623

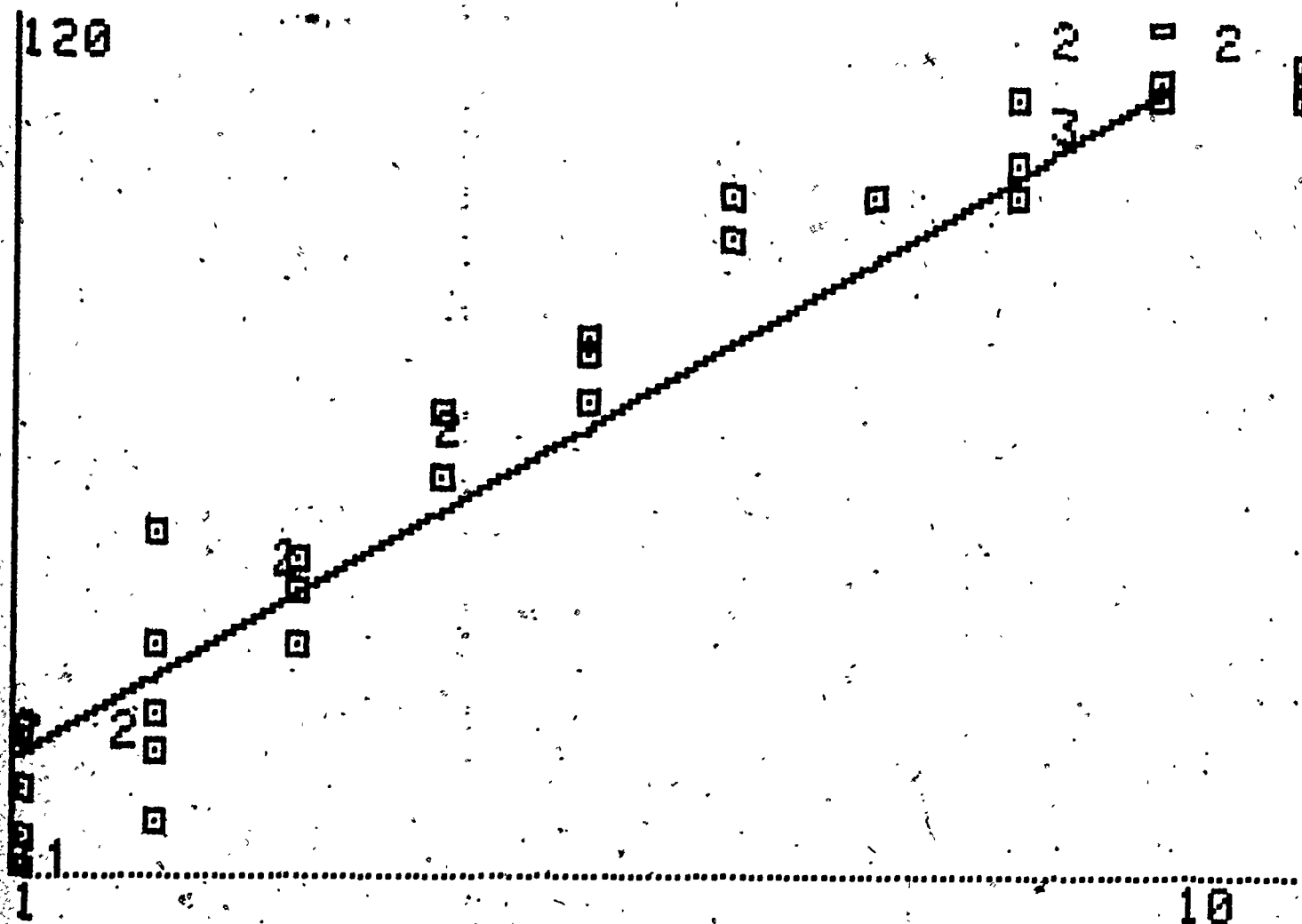
COEFFICIENT OF DETERMINATION (R SQR)= .9
36587579
COEFFICIENT OF CORRELATION= .96777455
STANDARD ERROR OF THE ESTIMATE= 9.193587
71

CARE TO LOOK AT THE BEST FIT? YES■

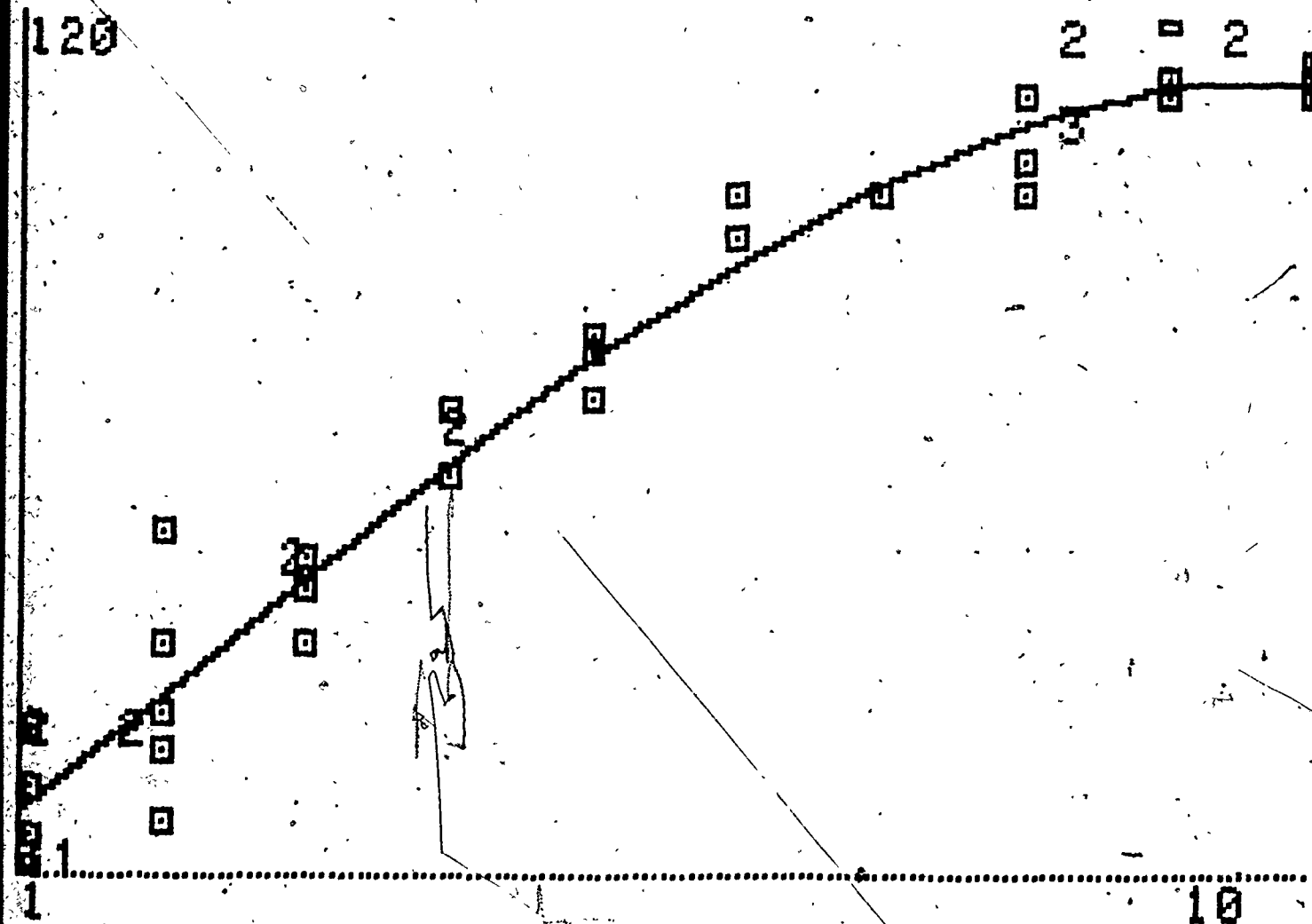
38-



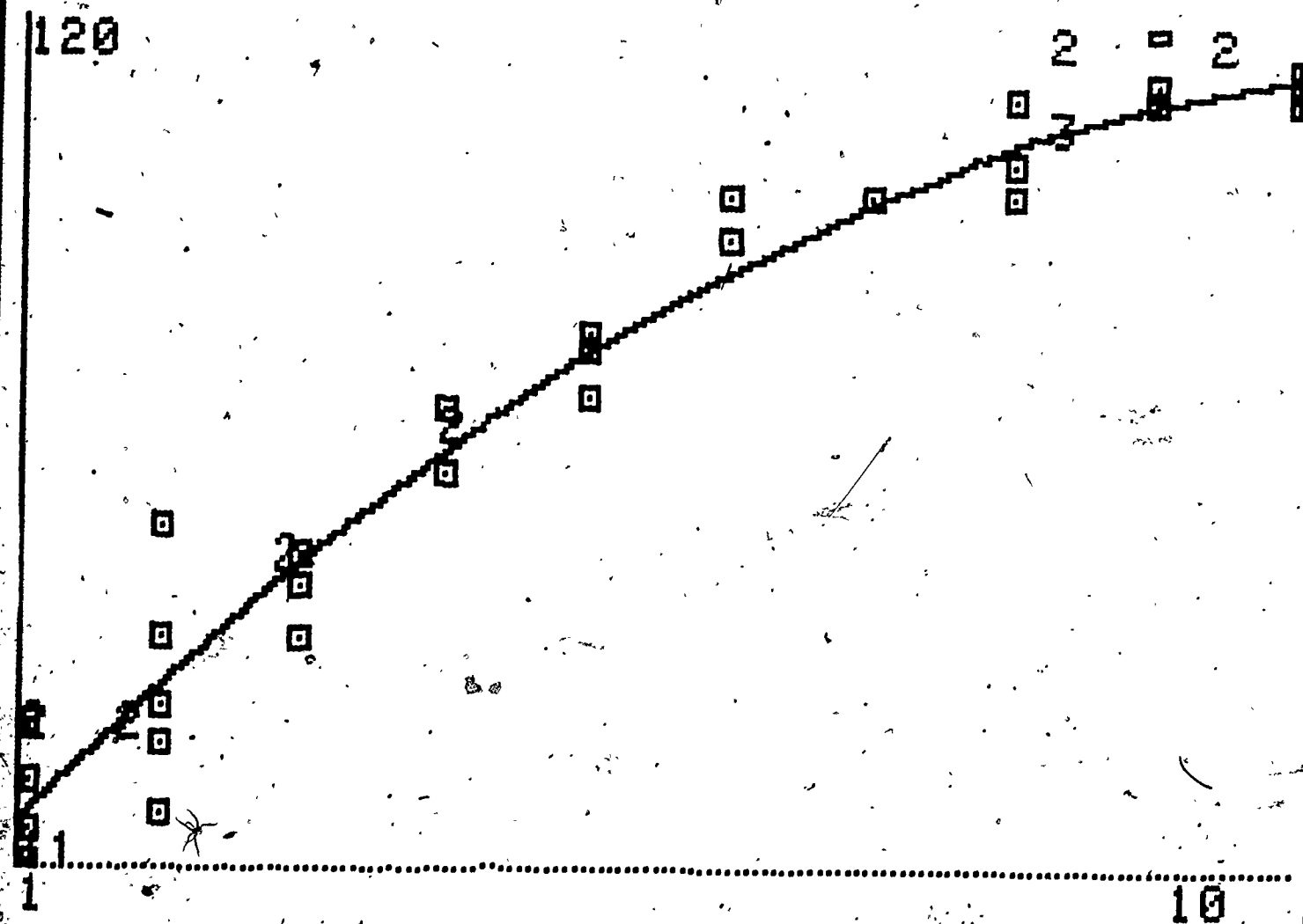
DEGREE OF THE EQUATION DESIRED



ANOTHER FIT? YES

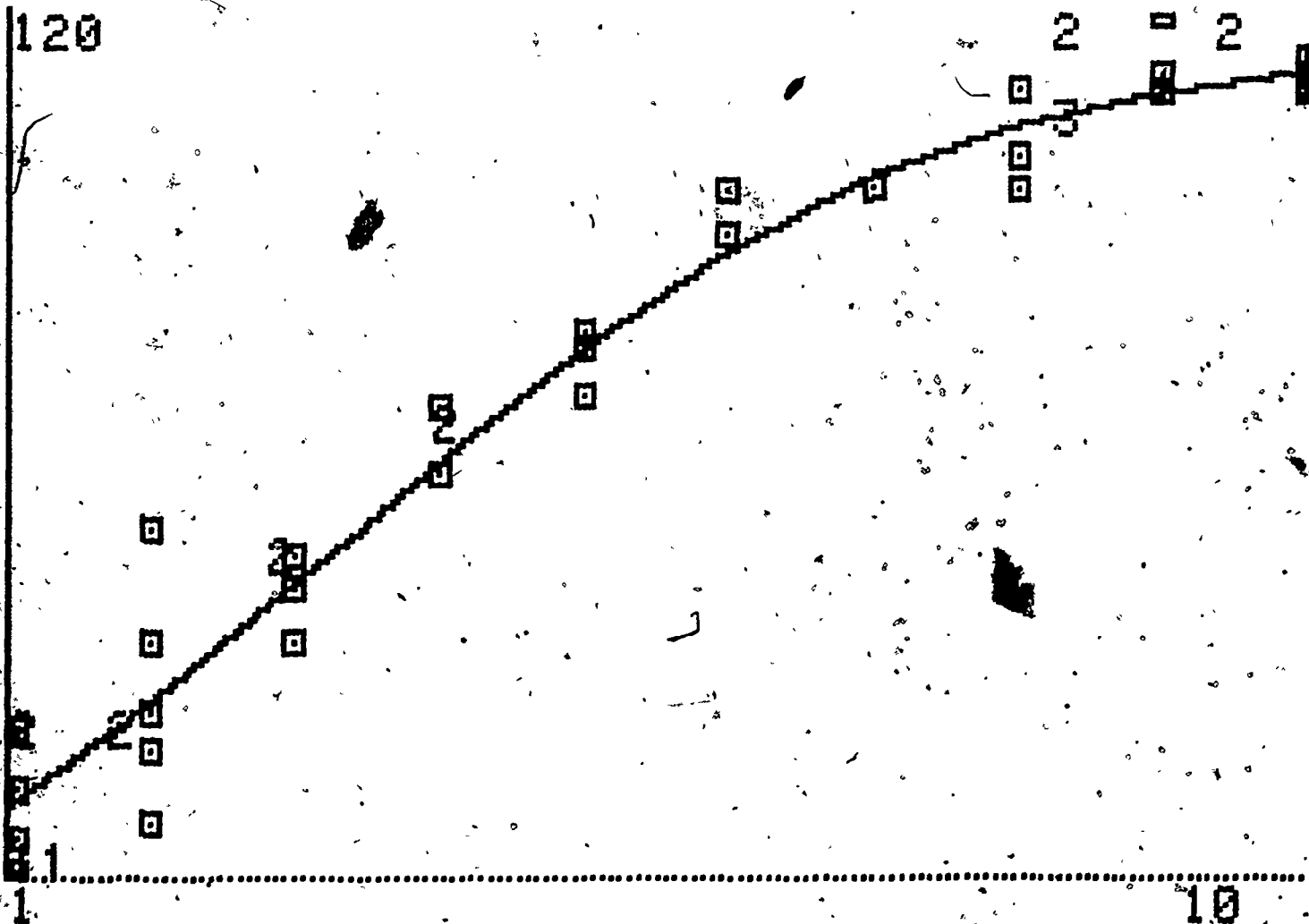


ANOTHER FIT? YES

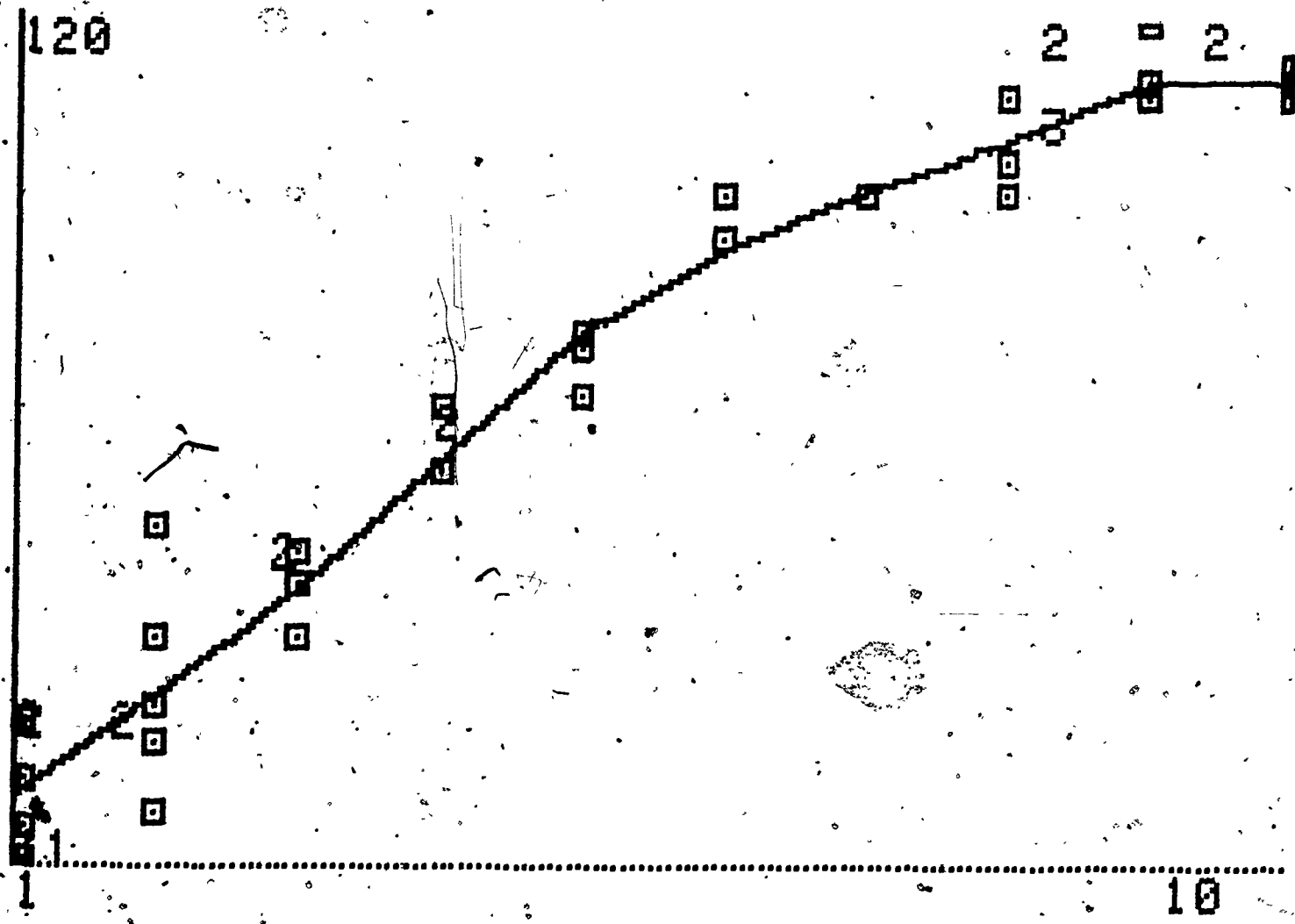


ANOTHER FIT? YES

48



ANOTHER FIT? YES



ANOTHER FIT? ■

1
 JRUN
 REGRESSION ANALYSIS
 NEED INSTRUCTIONS? YES
 INSERT THE NUMBER OF REPLICATES AT LINE20
 THEN TIMES OR PROBLEMS AT LINE 30
 THE CONTRASTS FOR TREND AT LINES 40 AND
 AT LINE160

20 N = 10
 25 REM REPLICATES/SUBJECTS
 30 P = 3: REM TIMES/PROBLEMS
 35 REM TIMES/PROBLEMS
 40 TT = 2: REM CONTRASTS FOR TRE
 ND
 45 REM CONTRASTS FOR TREND

160 DATA 1,1,0,-2,-1,1

TYPE CONT TO GO ON

BREAK IN 9047

JCONT

SAMPLE SIZE IS 30

1 1 1 0 0 0 0 0 0 0 0 50
 1 1 0 1 0 0 0 0 0 0 0 62.2
 1 1 0 0 1 0 0 0 0 0 0 101.1
 1 1 0 0 0 1 0 0 0 0 0 88
 1 1 0 0 0 0 1 0 0 0 0 114.7
 1 1 0 0 0 0 0 1 0 0 0 85.9
 1 1 0 0 0 0 0 0 1 0 0 62.4
 1 1 0 0 0 0 0 0 0 1 0 63.5
 1 1 0 0 0 0 0 0 0 0 1 88.6
 1 1 0 0 0 0 0 0 0 0 0 115
 0 -2 1 0 0 0 0 0 0 0 0 101.6
 0 -2 0 1 0 0 0 0 0 0 0 89.6
 0 -2 0 0 1 0 0 0 0 0 0 85.3
 0 -2 0 0 0 1 0 0 0 0 0 106.2
 0 -2 0 0 0 0 1 0 0 0 0 112.8
 0 -2 0 0 0 0 0 1 0 0 0 84
 0 -2 0 0 0 0 0 0 1 0 0 73.4
 0 -2 0 0 0 0 0 0 0 1 0 68
 0 -2 0 0 0 0 0 0 0 0 1 108
 0 -2 0 0 0 0 0 0 0 0 0 120
 -1 1 1 0 0 0 0 0 0 0 0 61.6
 -1 1 0 1 0 0 0 0 0 0 0 91.1
 -1 1 0 0 1 0 0 0 0 0 0 78.8
 -1 1 0 0 0 1 0 0 0 0 0 91.4
 -1 1 0 0 0 0 1 0 0 0 0 94
 -1 1 0 0 0 0 0 1 0 0 0 80.2
 -1 1 0 0 0 0 0 0 1 0 0 72.6
 -1 1 0 0 0 0 0 0 0 1 0 64
 -1 1 0 0 0 0 0 0 0 0 1 63.2
 -1 1 0 0 0 0 0 0 0 0 0 94

EQUATION COEFFICIENTS:

CONSTANT: 109.666667
 VARIABLE(1): 2.02500001
 VARIABLE(2): -4.59166667

VARIABLE(3): -38.6000001
 VARIABLE(4): -28.7000001
 VARIABLE(5): -21.2666667
 VARIABLE(6): -14.4666668
 VARIABLE(7): -2.5000009
 VARIABLE(8): -26.3000001
 VARIABLE(9): -40.2000001
 VARIABLE(10): -44.5000001
 VARIABLE(11): -23.0666667

BREAK IN 947
 JCONT

COEFFICIENT OF DETERMINATION(R-SQR)=.725080401

COEFFICIENT OF MULTIPLE CORRELATION =.85151653

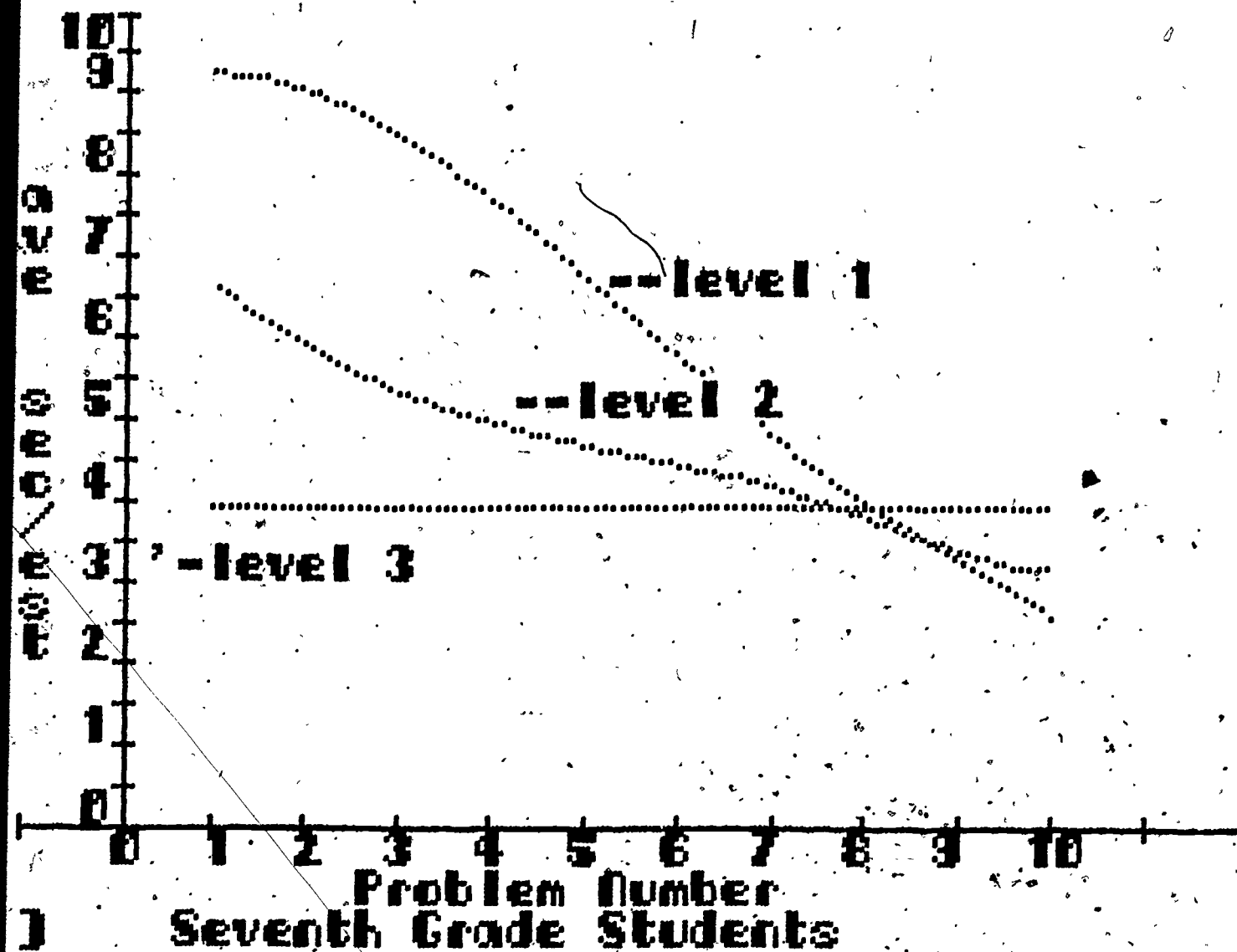
STANDARD ERROR OF ESTIMATE 14.27963

F VALUE = 4.31578981
 WITH 11 AND 18 DEGREES OF FREEDOM
 SIGNIFICANCE LEVEL IS 3.4E-03

SINGLE ORDER CORRELATIONS?YES

VAR#	R-SQR	F	SIG. LEVEL
1	7.8981211E-03	.5123	.5123
2	.121824795	7.97631861	.0109
3	.0688025652	4.50475768	.0456
4	7.21238924E-03	.5071	.5071
5	2.32864186E-03	.7021	.7021
6	.0289307665	1.89420398	.183
7	.147836582	9.67940625	6.1E-03
8	1.7577382E-03	.7373	.7373
9	.0846631703	5.54320999	.0285
10	.135432644	8.86727466	8E-03
11	2.56182574E-04	.8937	.8937

TOTAL R SQR IS .606943597



REFERENCES AND PROGRAMS

Kerlinger, F.N. and Pedhazur, E.J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart, and Winston, Inc. 1973.

Poole, L., and Borchers, M. Some Common Basic Programs. Berkeley: Osborne/McGraw-Hill, 1978.

NAME..NTH-CORKP DATE..03/29/81 TIME..850 PAGE..01
 STARTING LINE..0 ENDING LINE..63999

```

4  GOSUB 50000
5  HOME : PRINT :
10 PRINT : PRINT "CORRELATION PLOTTING"
50 PRINT "DO YOU NEED INSTRUCTIONS? "
60 INPUT Z$
70 IF LEFT$(Z$,1) = "Y" THEN 480
80 GOTO 10000
480 HOME : LIST 490 - 999: END
490 REM *** INSTRUCTIONS ***
500 REM TYPE IN DATA AS 510 DATA X1,Y1,...,999,999
502 REM WHERE X,Y ARE X VALUES, Y VALUES.
504 REM USE 999,999,999 TO STOP
510 DATA 2,4,2,6,2,5,4,7,4,10,4,10,6,13, 6,14,6,15,8,16,8,17,8,21,10,1
      8,10,19,10,20,12,19,12,20,12,21
999 DATA 999,999,999
10000 PRINT CHR$(4)"BRUN LOMEM:" & LOMEM: 24576
10010 HGR2 : PRINT CHR$(4)"BRUN AS CHR GEN"
10020 HOME
10030 DEF FN R(X) = INT (X * 1000 + .5) / 1000
10080 REM *** TO FIND MAX AND MIN TO FILL SCREEN ***
10085 CLEAR
10090 READ X,Y
10100 XI = X:XA = X:YI = Y:YA = Y
10110 READ X,Y
10120 IF X = 999 GOTO 10220
10130 IF X > XI GOTO 10150
10140 XI = X
10150 IF X < XA GOTO 10170
10160 XA = X
10170 IF Y > YI GOTO 10190
10180 YI = Y
10190 IF Y < YA GOTO 10110
10200 YA = Y
10210 GOTO 10110
10220 RESTORE
10230 HOME
10240 XS = 279 / (XA - XI):YS = 159 / (YA - YI)
10250 GOSUB 11170
10260 REM *** NTH-ORDER SUBROUTINE ***
10270 RESTORE
10280 VTAB 22
10290 INPUT "DEGREE OF THE EQUATION DESIRED";D
10300 HGR2
10305 PRINT "THIS MAY TAKE A LITTLE TIME....HMM.."
10310 DIM A(2 * D + 1),R(D + 1,D + 2),T(D + 2)
10320 REM ENTER DATA AND POPULATE MATRIX
10330 READ X
10340 IF X = 999 THEN 10480
10350 READ Y
10370 N = N + 1
10380 FOR J = 2 TO 2 * D + 1
10390 A(J) = A(J) + X ↑ (J - 1)
10400 NEXT J
10410 FOR K = 1 TO D + 1
10420 R(K,D + 2) = T(K) + Y * X ↑ (K - 1)
10430 T(K) = T(K) + Y * X ↑ (K - 1)
10440 NEXT K

```

NAME..NTH-CORKEP DATE..03/29/81 TIME..850 PAGE..02
STARTING LINE..0 ENDING LINE..63999

```

10450 T(D + 2) = T(D + 2) + Y * Y
10470 GOTO 10330
10480 A(1) = N
10490 FOR J = 1 TO D + 1
10500 FOR K = 1 TO D + 1
10510 R(J,K) = A(J + K - 1)
10520 NEXT K
10530 NEXT J
10540 FOR J = 1 TO D + 1
10550 FOR K = J TO D + 1
10560 IF R(K,J) < > 0 THEN 10600
10570 NEXT K
10580 PRINT "NON UNIQUE SOLUTION, SORRY!"
10590 GOTO 11090
10600 FOR I = 1 TO D + 2
10610 S = R(J,I)
10620 R(J,I) = R(K,I)
10630 R(K,I) = S
10640 NEXT I
10650 Z = 1 / R(J,J)
10660 FOR I = 1 TO D + 2
10670 R(J,I) = Z * R(J,I)
10680 NEXT I
10690 FOR K = 1 TO D + 1
10695 PRINT "STILL THINKING!"
10700 IF K = J THEN 10750
10710 Z = - R(K,J)
10720 FOR I = 1 TO D + 2
10730 R(K,I) = R(K,I) + Z * R(J,I)
10740 NEXT I
10750 NEXT K
10760 NEXT J
10765 HGR2
10770 VTAB 1: PRINT
10780 PRINT "CONSTANT= "; R(1,D + 2)
10790 FOR J = 1 TO D
10800 PRINT J; " DEGREE COEFFICIENT= "; R(J + 1,D + 2)
10810 NEXT J
10820 PRINT
10830 P = 0
10840 FOR J = 2 TO D + 1
10850 P = P + R(J,D + 2) * (T(J) - A(J) * T(1) / N)
10860 NEXT J
10870 Q = T(D + 2) - T(1) * 2 / N
10880 Z = Q - P
10890 I = N - D - 1
10900 J = P / D
10910 PRINT
10920 J = P / Q
10930 PRINT "COEFFICIENT OF DETERMINATION (R SQR.)= "; J
10940 PRINT "COEFFICIENT OF CORRELATION= "; SQR(J)
10950 PRINT "STANDARD ERROR OF THE ESTIMATE= "; SQR(ABS(Z / I))
10960 PRINT
10970 INPUT "CARE TO LOOK AT THE BEST FIT? "; Z$
10980 IF LEFT$(Z$,1) < > "Y" THEN 11090
10990 GOSUB 11170
11000 FOR X = XI TO XA

```

NAME..NTH-CORKP DATE..03/29/81 TIME..850 PAGE..03
STARTING LINE..0 ENDING LINE..63999

```

11010 P = R(1,D + 2)
11020 FOR J = 1 TO D
11030 P = P + R(J + 1,D + 2) * X + J
11040 NEXT J
11050 Y = P
11060 GOSUB 11110
11070 NEXT X
11080 VTAB 23
11090 INPUT "ANOTHER FIT? ";Z$
11095 IF LEFT$(Z$,1) = "Y" THEN 10085
11096 END
11100 LIST 840
11110 PX = INT (XS * (X - XI)):PY = 159 - INT (YS * (Y - YI))
11120 HCOLOR= 3
11130 IF PY < 1 THEN 11160
11140 IF PY > 159 THEN 11160
11150 HPLLOT EX,EY TO PX,PY
11155 EX = PX:EY = PY
11160 RETURN
11170 REM *** PLOT ROUTINE ***
11180 HGR2: HCOLOR= 2
11190 VTAB 1: PRINT YA: VTAB 20: HTAB 1: PRINT YI: HTAB 1: PRINT XI: HTAB
37: PRINT XA
11200 HPLLOT 0,0 TO 0,159 TO 279,159
11210 RESTORE
11220 READ X,Y
11230 IF X = 999 THEN RETURN
11250 PX = INT (XS * (X - XI)):PY = 159 - INT (YS * (Y - YI))
11260 HCOLOR= 3
11270 IF PX = 0 THEN PX = PX + 2
11280 IF PY = 159 THEN PY = PY - 1
11290 HPLLOT PX,PY
11300 GOSUB 11390
11320 GOTO 11220
11330 CX = INT (PX / 6.4):CY = INT (PY / 8)
11340 IF CX > 38 THEN CX = 38
11350 IF CX = 0 THEN CX = 1
11360 IF CY = 0 THEN CY = 1
11370 HTAB (CX): VTAB (CY): PRINT F
11380 GOTO 11220
11390 REM *** SQUARE SUBROUTINE ***
11400 XT = PX - 2:XB = PX + 2:YT = PY - 2:YB = PY + 2
11410 IF XT < 1 THEN XT = 1
11420 IF XB > 279 THEN XB = 279
11430 IF YT < 1 THEN YT = 1
11440 IF YB > 159 THEN YB = 159
11450 HPLLOT XT,YT TO XT,YB TO XB,YB TO XB,YT TO XT,YT
11460 RETURN
50000 HOME: VTAB 6: PRINT "::::::::::::::::::::::::::::::::::::::::"
50010 PRINT "::::::::::::::::::::::::::::::::::::::::"
50020 PRINT ";;"
50030 PRINT ";; ";; INVERSE: PRINT "-FITTING HIGHER ORDER PLOTS- ";; NORMAL
PRINT ";;"
50035 PRINT ";;"
50040 PRINT ";; CONTAINS SOME COMMON BASIC SUB"
50050 PRINT ";; PROGRAMS BY"
50060 PRINT ";;"

```

NAME..NTH-CORKP DATE..03/29/81 TIME..850 PAGE..04
STARTING LINE..0 ENDING LINE..63999

50070 PRINT "": FOR THE APPLE BY "":
50080 PRINT "": CARL F. BERGER "":
50090 PRINT "": UNIVERSITY OF MICHIGAN "":
50100 PRINT "": "":
50110 PRINT "::":
50120 PRINT "::":
50130 VTAB 22: INVERSE : PRINT "PRESS <RETURN> TO START": NORMAL
50140 GET Z\$: IF Z\$ = " " THEN 50140
50150 RETURN

** THE LINE CROSS REFERENCE **

PROGRAM NAME...NTH-COR
DATE...03/15/81

480: 10070
490: 480
10000: 10
10085: 11095
10110: 10190 10210
10150: 10130
10170: 10150
10190: 10170
10220: 10120
10330: 10470
10480: 10340
10600: 10560
10750: 10700
11090: 10590 10980
11110: 11060
11160: 11130 11140
11170: 10250 10990
11220: 11320 11380
11390: 11300

** THE VARIABLE CROSS REFERENCE **

PROGRAM NAME...NTH-COR
DATE...03/15/81

AC
10310 10390 10390 10480 10510 10850

CT
10360 10460

CX
11330 11340 11340 11350 11350 11370

CY
11330 11360 11360 11370

D
10290 10310 10310 10310 10310 10380 10410 10420 10450 10450 10490 10500 10540
10550 10600 10660 10690 10720 10780 10790 10800 10840 10850 10870 10890 10900
11010 11020 11030

EX
11150 11155

EY
11150 11155

F
10090 10110 10350 10360 11220 11240 11320 11370

I
10600 10610 10620 10620 10630 10640 10660 10670 10670 10680 10720 10730 10770

10730 10740 10890 10950 11240 11310

J

10380 10390 10390 10390 10400 10490 10510 10510 10530 10540 10550 10560 10610
10620 10650 10650 10670 10670 10700 10710 10730 10760 10790 10800 10800 10810
10840 10850 10850 10850 10860 10900 10920 10930 10940 11020 11030 11030 11040

K

10410 10420 10420 10420 10430 10430 10430 10440 10500 10510 10510 10520 10550
10560 10570 10620 10630 10690 10700 10710 10730 10730 10750

N

10370 10370 10480 10850 10870 10890

P

10830 10850 10850 10880 10900 10920 11010 11030 11030 11050

PX

11110 11150 11155 11250 11270 11270 11270 11290 11330 11400 11400

PY

11110 11130 11140 11150 11155 11250 11280 11280 11280 11290 11330 11400 11400

Q

10870 10880 10920

R

10030 10310 10420 10510 10560 10610 10620 10620 10630 10650 10670 10670 10710
10730 10730 10730 10780 10800 10850 11010 11030

S

10610 10630

T

10310 10420 10430 10430 10450 10450 10850 10850 10870 10870

X

10030 10030 10090 10100 10100 10110 10120 10130 10140 10150 10160 10330 10340
10390 10420 10430 11000 11030 11070 11110 11220 11230 11250

XA

10100 10150 10160 10240 11000 11190

XB

11400 11420 11420 11450 11450

XI

10100 10130 10140 10240 11000 11110 11190 11250

XS

10240 11110 11250

XT

11400 11410 11410 11450 11450 11450

Y

10090 10100 10100 10110 10170 10180 10190 10200 10350 10420 10430 10450 10450
11050 11110 11220 11250

YA

10100 10190 10200 10240 11190

YB

11400 11440 11440 11450 11450

YI

10100 10170 10180 10240 11110 11190 11250

YS

10240 11110 11250

YT

11400 11430 11430 11450 11450 11450

Z

10650 10670 10710 10730 10880 10950

Z\$

10060 10070 10970 10980 11090 11095